

RAW SEQUENCE LISTING DATE: 05/08/2001
 PATENT APPLICATION: US/09/430,590E TIME: 14:41:23

Input Set : A:\ES.txt
 Output Set: N:\CRF3\05082001\I430590E.raw

78 <222> LOCATION: (398)..(1372)
 79 <223> OTHER INFORMATION: ORF1 coding sequence for gag
 82 <220> FEATURE:
 83 <221> NAME/KEY: CDS
 84 <222> LOCATION: (1373)..(6103)
 85 <223> OTHER INFORMATION: ORF2 - coding sequence for pol
 88 <400> SEQUENCE: 3

89	tgtggtttg	tgcactattt	tgtgtcagaa	actgatcaat	gaaaatgatg	gttattatga	60											
91	aatggaaaa	ttttccatc	acacatcagg	tgtgacaga	actaaactat	attgtgtagt	120											
93	ataaataagg	gtatgaaata	ccaacatccc	agaatatcaa	cgagatagaa	gggaggagtt	180											
95	tcaatatata	tcttgtgaat	aataacttcg	ttcttaattca	ctatacacaa	ctagacgtgt	240											
97	acacgctcaa	tctcaggtaa	agaaagttt	tattccatca	gattagaagt	cgatagtgtat	300											
99	aatcatttcg	tcccaaattt	gcgttgtata	aattcagtc	tcagattgt	attattgatt	360											
101	gatagtttcg	aagtttgaag	gtacagaatt	tcacaag	atg	agt	6103											
102				tcc	gca	aag	415											
103				Met	Ser	Ser	Ala	Lys	Asn									
104				1		5												
105	gat	gat	aac	gaa	ggg	aag	gtc	atg	gaa	agt	gtt	gat	caa	gct	aat	gct	463	
106	Asp	Asp	Asn	Glu	Gly	Lys	Val	Met	Glu	Ser	Val	Asp	Gln	Ala	Asn	Ala		
107	10						15				20							
108	att	agt	aag	gtg	gat	gaa	cat	atc	aag	gct	aga	ttc	aat	atg	ctt	ttc	511	
109	Ile	Ser	Lys	Val	Asp	Glu	His	Ile	Lys	Ala	Arg	Phe	Asn	Met	Leu	Phe		
110	25						30				35							
111	ata	aaa	ttt	aat	gac	tta	cct	aag	ttg	gcc	gtc	ggt	aat	cag	aaa	agc	559	
112	Ile	Lys	Phe	Asn	Asp	Leu	Pro	Lys	Leu	Ala	Val	Gly	Asn	Gln	Lys	Ser		
113	40						45				50							
114	gtg	gat	aaa	tgg	aat	gaa	ttt	aaa	tat	ttc	cac	gtt	gct	tac	ccc	607		
115	Val	Asp	Lys	Trp	Asn	Glu	Glu	Phe	Lys	Tyr	Phe	His	Val	Ala	Tyr	Pro		
116	55						60				65			70				
117	gat	gtt	ttg	gaa	ttt	ttg	ctt	gac	tat	aat	cct	aaa	gat	aaa	ttc	aag	655	
118	Asp	Val	Leu	Glu	Leu	Asp	Tyr	Asn	Pro	Lys	Asp	Phe	Lys	Asp	Phe	Lys		
119	121	75					80				85							
120	gtt	aaa	aag	gta	gaa	ggt	att	tat	ttt	act	ggt	tgg	tgt	tta	caa	atg	703	
121	Val	Lys	Val	Glu	Gly	Ile	Tyr	Phe	Thr	Gly	Trp	Cys	Leu	Gln	Met			
122	90						95				100							
123	125	105					110				115							
124	tgt	tta	cag	tcc	att	ttt	gat	agg	ttc	aga	ttg	atc	atg	att	tct	aag	751	
125	Cys	Leu	Gln	Ser	Ile	Phe	Asp	Arg	Phe	Arg	Leu	Ile	Met	Ile	Ser	Lys		
126	131	120	105				110				115							
127	cta	cca	aag	cac	ttg	caa	aag	gaa	gca	aac	tta	atc	aaa	gct	gct	tat	799	
128	Leu	Pro	Lys	His	Leu	Gln	Lys	Glu	Ala	Asn	Leu	Ile	Lys	Ala	Ala	Tyr		
129	133	120	105				125				130							
130	gat	gct	gtt	act	aaa	tct	aaa	gat	tat	acc	att	act	agt	aag	atc	ttg	847	
131	Asp	Ala	Val	Thr	Lys	Ser	Lys	Asp	Tyr	Thr	Ile	Thr	Ser	Lys	Ile	Leu		
132	135	120	105				140				145							
133	ctg	aag	ttt	gta	aac	gtt	gaa	cat	gag	tta	gtg	gtt	tgc	tat	aac	ctt	895	
134	Ser	Lys	Phe	Val	Asn	Val	Glu	His	Glu	Leu	Val	Val	Cys	Tyr	Asn	Leu		
135	141	120	105				155				160			165				
136	cca	tat	ttg	ctg	cag	gtg	gaa	gag	aaa	ctt	gag	gaa	ata	ctc	tac	aac	943	
137	Pro	Tyr	Leu	Ser	Gln	Val	Glu	Glu	Lys	Leu	Glu	Glu	Ile	Leu	Tyr	Asn		
138	147	120	105				170				175			180				

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/430,590E

DATE: 05/08/2001

TIME: 14:41:23

Input Set : A:\ES.txt

Output Set: N:\CRF3\05082001\I430590E.raw

149	act tca aac gtt gtc gat gag tat gtc cgt agt ctt cca aat ctc ata	991
150	Thr Ser Asn Val Val Asp Glu Tyr Val Arg Ser Leu Pro Asn Leu Ile	
151	185 190 195	
153	ggt caa gtc ttg tac ttc aat cat gtg aag aaa tca gag gct tta agt	1039
154	Gly Gln Val Leu Tyr Phe Asn His Val Lys Lys Ser Glu Ala Leu Ser	
155	200 205 210	
157	ttg ttt ttg aat att cat gcc tca tac tac tca aag tgg att caa gct	1087
158	Leu Phe Leu Asn Ile His Ala Ser Tyr Tyr Ser Lys Trp Ile Gln Ala	
159	215 220 225 230	
161	gac aat gat aca tca gta ctc cca agt tgc tct acc ata gct gaa gaa	1135
162	Asp Asn Asp Thr Ser Val Leu Pro Ser Cys Ser Thr Ile Ala Glu Glu	
163	235 240 245	
165	atg tgt gat cat cct gat tat gct aga ttg gtt gac att cca agc aac	1183
166	Met Cys Asp His Pro Asp Tyr Ala Arg Leu Val Asp Ile Pro Ser Asn	
167	250 255 260	
169	aaa tat gaa ctt aat ctt att gtt agt tta cca gca cca gag aaa cca	1231
170	Lys Tyr Glu Leu Asn Leu Ile Val Ser Leu Pro Ala Pro Glu Lys Pro	
171	265 270 275	
173	aaa gga aaa cca gag gag aac tca ctg gaa caa tct caa aag aag aac	1279
174	Lys Gly Lys Pro Glu Glu Asn Ser Ser Glu Gln Ser Gln Lys Lys Asn	
175	280 285 290	
177	ctg aaa tca aga aag aga aat aag aaa cat cca aaa tca gat aac gat	1327
178	Ser Lys Ser Arg Lys Arg Asn Lys Lys His Pro Lys Ser Asp Asn Asp	
179	295 300 305 310	
181	aaa ggt gaa aaa gaa aaa gaa aaa act tca ctg gaa tga aaa	1375
182	Lys Gly Glu Lys Glu Lys Glu Lys Lys Thr Ser Ser Glu Lys	
183	315 320 325	
185	aca ggt gct tct att aat tgt gta atg aat ata cat aat tgc agc	1423
186	Thr Gly Ala Ala Ser Ile Asn Cys Val Met Asn Ile His Asn Cys Ser	
187	330 335 340	
189	aaa acc acg ttt cca gta gaa aat tct cat tct ctt aat gct tct ttg	1471
190	Lys Thr Thr Pro Val Glu Asn Ser His Ser Leu Asn Ala Ser Leu	
191	345 350 355	
193	aac gta atg aat ttt aaa ggt tta agg ttt aac aag tat cta gtg tat	1519
194	Asn Val Met Asn Phe Lys Gly Leu Arg Phe Asn Lys Tyr Leu Val Tyr	
195	360 365 370	
197	gat act ggt gcc aca ata tct gtt gtg aac aat aaa gat ata ttg ctg	1567
198	Asp Thr Gly Ala Thr Ile Ser Val Val Asn Asn Lys Asp Ile Leu Ser	
199	375 380 385	
201	aat gtt aag gac gca aca att gaa gtt tct gtt gct gat ggt gct aca	1615
202	Asn Val Lys Asp Ala Thr Ile Glu Val Ser Val Ala Asp Gly Ala Thr	
203	390 395 400 405	
205	tta gaa gca gat tgt att ggt gat cta att atc aga gtc ggt att gtc	1663
206	Leu Glu Ala Asp Cys Ile Gly Asp Leu Ile Ile Arg Val Gly Ile Val	
207	410 415 420	
209	tcg att acg tta gag aat aca ttg tat tta cca gaa agt tcc ttt aat	1711
210	Ser Ile Thr Leu Glu Asn Thr Leu Tyr Leu Pro Glu Ser Ser Phe Asn	
211	425 430 435	
213	ctt gtg agt ttg aaa caa att gaa gaa cga gga ttt aat gtt ctt att	1759

RAW SEQUENCE LISTING DATE: 05/08/2001
PATENT APPLICATION: US/09/430,590E TIME: 14:41:23

Input Set : A:\ES.txt
Output Set: N:\CRE3\05082001\I430590E.raw

214	Leu	Val	Ser	Leu	Lys	Gln	Ile	Glu	Glu	Arg	Gly	Phe	Asn	Val	Leu	Ile	
215				440				445				450					
217	act	aaa	gaa	tca	gtg	att	gta	ttt	aac	caa	aat	gtg	gct	cct	act	att	
218	Thr	Lys	Glu	Ser	Val	Ile	Val		Phe	Asn	Gln	Asn	Val	Ala	Pro	Thr	Ile
219				455				460				465					
221	att	gct	tca	agg	aag	aat	gct	gct	gat	ctt	tat	atg	ggg	cct	caa	ttc	
222	Ile	Ala	Ser	Arg	Lys	Asn	Ala	Ala	Asp	Leu	Tyr	Met	Gly	Pro	Gln	Phe	
223	470						475				480				485		
225	agt	gaa	gaa	tct	tta	gaa	tgt	gat	ttt	gat	tat	gat	ggg	ttg	gca	gat	
226	Ser	Glu	Glu	Ser	Leu	Glu	Cys	Asp	Phe	Asp	Tyr	Asp	Gly	Leu	Ala	Asp	
227							490			495				500			
229	atg	ttg	tcc	aat	gct	aac	caa	gat	gac	aaa	gat	aaa	tca	agt	atg	aat	
230	Met	Leu	Ser	Asn	Ala	Asn	Gln	Asp	Asp	Lys	Asp	Lys	Ser	Ser	Met	Asn	
231							505			510				515			
233	gaa	atg	tca	gaa	tat	caa	gaa	cat	gat	tat	agt	tct	cga	gca	tta	ata	
234	Glu	Met	Ser	Glu	Tyr	Gln	Glu	His	Asp	Tyr	Ser	Ser	Arg	Ala	Leu	Ile	
235							520			525				530			
237	aat	tct	ttg	acg	gag	gtt	gat	gtt	tta	gat	gtt	gaa	att	tcc	cca	tat	
238	Asn	Ser	Leu	Thr	Glu	Val	Asp	Val	Leu	Asp	Val	Glu	Ile	Ser	Pro	Tyr	
239							535			540				545			
241	gga	gtt	gaa	caa	ttg	cta	cca	act	gga	gat	aag	aac	gat	att	tat	aat	
242	Gly	Val	Glu	Gln	Leu	Leu	Pro	Thr	Gly	Asp	Lys	Asn	Asp	Ile	Tyr	Asn	
243	550						555				560				565		
245	ttc	cat	ttg	atg	tca	aat	cat	atg	tcc	att	gag	aaa	atc	ttg	ttg	tta	
246	Phe	His	Leu	Met	Ser	Asn	His	Met	Ser	Ile	Glu	Lys	Ile	Leu	Leu		
247							570			575				580			
249	caa	aaa	tac	cag	ggt	ctc	gta	ctt	cac	act	tca	aaa	gag	agt	ctt	caa	
250	Gln	Lys	Tyr	Gln	Gly	Leu	Val	Leu	His	Thr	Ser	Lys	Glu	Ser	Leu	Gln	
251							585			590				595			
253	aag	att	gct	gat	tgt	aag	gta	tgt	cta	tta	tcg	aat	gcc	aaa	cag	aga	
254	Lys	Ile	Ala	Asp	Cys	Lys	Val	Cys	Leu	Leu	Ser	Asn	Ala	Lys	Gln	Arg	
255							600			605				610			
257	agt	cac	aat	cat	cat	tca	gaa	aga	aaa	gcc	tcg	aga	aga	cat	gag	aga	
258	Ser	His	Asn	His	His	Ser	Glu	Arg	Lys	Ala	Ser	Arg	Arg	His	Glu	Arg	
259							615			620				625			
261	ctt	cat	tgt	gat	act	ctc	ggt	cca	ttt	agg	tcc	gaa	aat	aac	aag	tgg	
262	Leu	His	Cys	Asp	Thr	Leu	Gly	Pro	Phe	Arg	Ser	Glu	Asn	Asn	Lys	Trp	
263	630						635				640				645		
265	tat	tta	acg	tct	gtt	ata	gat	gaa	cat	acg	ggt	tac	att	gaa	gga	att	
266	Tyr	Leu	Thr	Ser	Val	Ile	Asp	Glu	His	Thr	Gly	Tyr	Ile	Glu	Gly	Ile	
267							650			655				660			
269	att	act	aaa	gac	aga	aag	gta	aag	gat	ctc	tta	att	caa	cga	tta	aag	
270	Ile	Thr	Lys	Asp	Arg	Lys	Val	Lys	Asp	Leu	Leu	Ile	Gln	Arg	Leu	Lys	
271							665			670				675			
273	atc	tgg	aat	aat	cgg	ttt	aac	gat	aag	gtg	gca	tac	ttc	aga	agt	gat	
274	Ile	Trp	Asn	Asn	Arg	Phe	Asn	Asp	Lys	Val	Ala	Tyr	Phe	Arg	Ser	Asp	
275							680			685				690			
277	aat	gct	cct	gag	ttc	cca	caa	cct	tct	gat	tta	gct	gag	ttc	ggt	att	
278	Asn	Ala	Pro	Glu	Phe	Pro	Gln	Pro	Ser	Asp	Leu	Ala	Glu	Phe	Gly	Ile	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/430,590E

DATE: 05/08/2001

TIME: 14:41:23

Input Set : A:\ES.txt

Output Set: N:\CRF3\05082001\I430590E.raw

279	695	700	705		
281	tgg	agg gag act ata	gcg gca tat ctg	cct gag ctt aat ggt ctc gcc	2575
282	Trp	Arg Glu Thr Ile	Ala Ala Tyr Ser Pro	Glu Leu Asn Gly Leu Ala	
283	710		715	720	725
285	gag	gtt aat aaa ttg	att tta caa cag	att tac agg atc gtt gtg	2623
286	Glu	Val Val Asn Lys	Leu Ile Leu Gln	Gln Ile Tyr Arg Ile Val Val	
287			730	735	740
289	aca	ctt ggt cca caa	ata ctc aag ttg	att tat tat gtg	2671
290	Thr	Leu Gly Pro Gln Ile	Leu Lys Leu Ile	Tyr Tyr Val Ile Gln Tyr	
291			745	750	755
293	tct	att aca atg atc aac	cac act cca cgt	cgt tca ctc aag gga caa	2719
294	Ser	Ile Thr Met Ile Asn His	Thr Pro Arg Arg	Ser Leu Lys Gly Gln	
295			760	765	770
297	acc	cct tat ggt tgc	tat tat caa tta	agt gag gga aat ttc tac	2767
298	Thr	Pro Tyr Gly Cys	Tyr Tyr Gln	Leu Ser Glu Gly Asn Phe	
299			775	780	785
301	ttt	cct ttt gcc atc	gat tgt gtc	gtt aca ttt agt aat gcc	2815
302	Phe	Pro Phe Ala Ile	Asp Cys Val Val	Thr Phe Ser Asn Ala Ile	
303	790		795	800	805
305	aag	aac cgt tac gga	gtt aca tca act	aaa gga gct cct tca tcg	2863
306	Lys	Asn Arg Tyr Gly	Val Thr Ser Thr	Lys Gly Ala Pro Ser Ser	
307			810	815	820
309	atg	ggg gct gtg	att ggc tac gct	agc gat tgt ttt agt tat tac	2911
310	Met	Gly Ala Val Ile	Gly Tyr Ala Ser	Asp Cys Phe Ser Tyr Tyr	
311			825	830	835
313	ttg	cta aaa aat atg	cgg tgt gat	att atc ctt agc cct aat gtc	2959
314	Leu	Leu Lys Asn Met	Arg Cys Asp Ile	Ile Leu Ser Pro Asn Val	
315			840	845	850
317	ata	ttg cga agc tat	gag gtt att aac	tcc tat ctc aaa aac tta tcc	3007
318	Ile	Leu Arg Ser Tyr	Glu Val Ile	Asn Ser Tyr Leu Lys Asn	
319			855	860	865
321	act	aca cct atg tca	cac att gtt cct	atg gct gaa ggt atc cag	3055
322	Thr	Thr Pro Met Ser His	Ile Val Pro Met	Ala Glu Gly Ile Gln Gly	
323	870		875	880	885
325	agg	caa ctg ggc	gct cag tac gag	gta cgc gga aca tat gtg	3103
326	Arg	Gln Ser Gly Ala	Gln Tyr Glu Val	Arg Gly Thr Tyr Val	
327			890	895	900
329	gaa	tat gac aat aca	aat gac gtg	atg cac atg ccc	3151
330	Glu	Tyr Asp Asn Thr	Asn Asp Val	Met His Met Pro Lys Glu	
331			905	910	915
333	tca	gtt cag cca	gca tcg ttt	act tta act acg	3199
334	Ser	Val Gln Pro Ala	Ser Phe Thr	Leu Thr Thr Gly	
335			920	925	930
337	gaa	tat gtt ata aat	gat gat cca	gta cag att acc	3247
338	Glu	Tyr Val Ile Asn	Asp Asp Pro	Val Gln Ile Thr	
339			935	940	945
341	gat	gat ttt tct aac	cct ctt caa	cta act gaa gaa	3295
342	Asp	Asp Phe Ser Asn	Pro Leu Gln	Leu Thr Glu	
343	950		955	960	965

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/430,590E

DATE: 05/08/2001
TIME: 14:41:24

Input Set : A:\ES.txt
Output Set: N:\CRF3\05082001\I430590E.raw

L:1161 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:1185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1187 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1189 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1191 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1197 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1199 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1201 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1203 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1217 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:1501 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:1613 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1617 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1729 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1731 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1935 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:2131 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:2255 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:2275 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:2323 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:2325 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:2327 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:2365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19
L:2505 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:3156 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:3184 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:3190 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:3192 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:3278 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:3280 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:3438 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28
L:3502 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28
L:3552 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28
L:3634 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29
L:3654 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3656 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3768 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:3770 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:3772 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:3774 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:3776 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:3778 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:3830 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32
L:3886 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33
L:4002 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33
L:4004 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33
L:5088 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/430,590E

DATE: 05/08/2001

TIME: 14:41:24

Input Set : A:\ES.txt

Output Set: N:\CRF3\05082001\I430590E.raw

L:6060 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:111

L:8916 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139

L: 9695 M: 258 W: Mandatory Feature missing, <220> FEATURE:

L: 9695 M: 258 W: Mandatory Feature missing, <220> FEATURE:
L: 9695 M: 258 W: Mandatory Feature missing, <223> OTHER INFORMATION: